Supplementary Material to “Dispersion in options investors’ versus analysts’ expectations: Predictive inference for stock returns”

Not for publication

August 2019

Abstract

This supplementary material includes additional results not presented in the main body of the paper.

A DISP with Signed Volume

In the main paper, DISP is estimated using raw options volume data. This is supported by the empirical evidence provided by Hu (2014) that end-users are typically net buyers of out-of-the-money (OTM) options and net sellers of in-the-money (ITM) options.

In this section, we compare the main DISP measure with a sophisticated DISP measure that utilizes only the end-users' buy-side volume of OTM options and the end-users' sell-side volume of ITM options (DISP_SIGNED). In essence, for the construction of DISP_SIGNED we retain only OTM call purchases and ITM put sales, which are undoubtedly optimistic trades related to positive expectations, and OTM put purchases and ITM call sales, which are undoubtedly pessimistic trades related to negative expectations. To do so, we obtain signed volume data from the International Securities Exchange (ISE) Trade Profile. This dataset contains all end-users' trades disaggregated by whether each trade is a buy or a sell order. While the ISE options volume data represent about 30% of the total individual stock options trading volume across all options exchanges, Ge et al. (2016) show that the data are quite representative of the total options volume provided by OptionMetrics. Unfortunately, the ISE data are only available from May 2005 onwards, thus limiting considerably the period for which we can estimate DISP.

Figure 1 depicts DISP (red solid line) and DISP_SIGNED (black dashed line) for the period 2005:05-2015:08, both standardized to have zero mean and variance equal to one. It can be seen that DISP behaves very similarly to DISP_SIGNED. In fact, the correlation between the two variables is 95%. The above evidence shows that a simple DISP measure that uses unsigned volume data captures almost exactly the same information with a sophisticated DISP measure that uses signed volume data. Therefore, given the considerably longer sample period covered by OptionMetrics than by ISE Trade Profile, it is natural that our analysis is conducted with unsigned volume data. It is also important to note that, unlike signed volume data, daily unsigned volume data are publicly available and hence easily accessible to investors. This means that the trading strategy based on the out-of-sample predictive power of DISP would be relatively cheap and implementable by an investor in real time.

B Alternative DISP Measures

In this section, we provide in-sample, out-of-sample and economic significance results for eight alternative dispersion measures. In particular, DISP_{I} and DISP_{II} are constructed using the volume-weighted mean absolute deviation and interquartile range of moneyness levels, respectively. DISP_{III} follows conceptually the dispersion in beliefs measure of Diether et al. (2002) and utilizes the volume-weighted standard deviation of the strike prices scaled by the volume-weighted average strike price. For DISP_{IV} we remove options with moneyness between 0.975 and 1.025, while for DISP_{V} we use all available moneyness levels. DISP_{VI} employs only the end-of-month DISP value for each stock rather than the average value of the given month. Finally, DISP_{VII} uses only options that expire on the next standard expiration date,[1] while $\text{DISP}_{\text{VIII}}$ uses only options that expire one month after the next available standard expiration date. This way our dispersion measure is estimated using always only options that have the same expiration date. The results provided in Tables 1 - 2 are very similar to those provided for the main DISP measure and show that the predictive power of the dispersion in options investors' expectations is robust to alternative specifications and filtering rules.

C Description of the Bootstrap Method

This section describes the wild bootstrap procedure for computing the empirical p-values. Similar procedures are also followed by Neely et al. (2014) and Huang et al. (2015).

We begin by estimating the error terms from a regression of the future market return on the set of predictors used in the study:

$$\widehat{\varepsilon}_{t,t+h} = re_{t,t+h} - \left(\widehat{\alpha}_h + \widehat{\boldsymbol{\beta}}_h' \mathbf{z}_t\right), \tag{1}$$

where $re_{t,t+h}$ is the h-month excess market return, $\mathbf{z}_t$ is the vector of predictors and $\widehat{\alpha}_h$ and $\widehat{\boldsymbol{\beta}}_h'$ are the estimated OLS parameters.

Following convention, each predictor i included in the model of equation (1) is assumed to follow

[1] A standard maturity option is one that expires on the third Friday of a given month.

an AR(1) process:

$$z_{i,t+1} = \rho_{i,0} + \rho_{i,1} z_{i,t} + \phi_{i,t+1}. \tag{2}$$

For each predictor i define:

$$\widehat{\phi}^{c}_{i,t+1} = z_{i,t+1} - \left(\widehat{\rho}^{c}_{i,0} + \widehat{\rho}^{c}_{i,1} z_{i,t}\right), \tag{3}$$

where $\widehat{\rho}^{c}_{i,0}$ and $\widehat{\rho}^{c}_{i,1}$ are reduced-bias estimates of the respective AR(1) parameters in (2). The reduced-bias estimates are computed by iterating on the analytical second-order bias expression for the OLS estimates.

Using these reduced-bias AR(1) parameters and the fitted error terms from (1) and (3) we build up a pseudo-sample for the excess market return under the null of no predictability and for each of the predictive variables:

$$\widetilde{re}_{t,t+h} = \overline{re}_{t,t+h} + \widehat{\varepsilon}_{t,t+h} w_{t+1}, \tag{4}$$

$$\widetilde{z}_{i,t+1} = \widehat{\rho}^{c}_{i,0} + \widehat{\rho}^{c}_{i,1} \widetilde{z}_{i,t} + \widehat{\phi}^{c}_{i,t+1} w_{t+1}, \tag{5}$$

where $\overline{re}_{t,t+h}$ is the sample mean of the market excess return, w_{t+1} is a draw for the standard normal distribution and $\widetilde{z}_{i,t}$ for $t = 0$ is the initial value $z_{i,0}$ for each predictor. By multiplying $\widehat{\varepsilon}_{t,t+h}$ and each predictor's $\widehat{\phi}^{c}_{i,t+1}$ with the same draw from the standard normal distribution w_{t+1}, we are able to account for the cross-correlation between the market returns and the innovations in the predictive variables (Stambaugh, 1999) as well as for general forms of conditional heteroskedasticity. In addition, the reduced-bias AR(1) estimates ensure that the high persistence of several predictive variables is properly captured.

Finally, for each regression model examined in the paper, we estimate the Newey-West t-statistics using the equity premium and appropriate predictor time-series from the constructed pseudo-sample. By repeating the process 2,000 times, we obtain an empirical distribution for each of the t-statistics. The empirical p-value for each predictor in each regression model is the proportion of the bootstrapped t-statistics that exceed in absolute terms the respective Newey-West t-statistic from the original sample.

D Alternative Market-Timing Strategies

In the main paper, we present the results from a market-timing strategy that utilizes mean-variance portfolio weights. In this section, we provide additional results from a strategy that utilizes binary weights. In particular, we consider two scenarios: one where short-sales are not allowed and one where short-sales are allowed. In the first scenario, the investor allocates 100% of her wealth in the market index (risk-free asset) every time the predicted equity premium is positive (negative). In the second scenario, the investor allocates 150% of her wealth in the market index (risk-free asset) and –50% in the risk-free asset (market index) every time the predicted equity premium is positive (negative).

The historical average (HAV) strategy delivers a Sharpe ratio ranging from 0.48 to 0.51 when short-sales are not allowed and from 0.45 to 0.48 when short-sales are allowed. It can be seen that DISP clearly outperforms the HAV strategy in all scenarios since it provides higher Sharpe ratios and positive ΔCER values. Moreover, unlike the case of the mean-variance strategy discussed in the main paper, in the case of binary strategies AFD performs worse than the historical average delivering lower Sharpe ratios and negative ΔCER values.

E Sample without the financial crisis

In this section, we examine whether the strong predictive power of DISP for the equity premium is driven by the financial crisis period. In particular, we consider the pre-crisis period (1996:01-2008:06), as well as the whole sample period excluding the financial crisis (1996:01-2008:06 and 2009:07-2017:12). Tables 4 and 5 provide the results, while Figure 2 plots the respective cumulative square prediction error differences.

Overall, DISP continues to be a strong predictor of future market returns even when the financial crisis period is not taken into consideration. One exception is the 1-month ahead predictability in the pre-crisis period where the results are somewhat weaker. However, a closer inspection of Figure 2 reveals that even in this case DISP consistently outperforms the historical average model across the majority of the months. The low statistical significance is driven by one observation, namely DISP in 1998:07 predicting the return of 1998:08. This observation corresponds to the outburst of the Russian financial crisis, when the US stock market experienced a monthly return of almost –16% (the second lowest in our sample after the return of 2008:10). By removing this one observation from the pre-crisis sample, DISP becomes significant at the 5% level with a t-stat of –2.33 and a p-value from the wild bootstrap experiment of 0.024.

F Time-series plots of DISP versus ASYM and AFD versus AFA

In the main paper, we report the correlations between the dispersion measures under examination (DISP and AFD) and the corresponding asymmetry measures (ASYM and AFA). Figure 3 compares the time-series pattern of DISP versus ASYM, and AFD versus AFA. It is evident that after 2007 AFD and AFA tend to move in opposite directions.

References

[1] Clark, T. and M. McCracken. 2001. Tests of equal forecast accuracy and encompassing for nested models. Journal of Econometrics 105, 85-110.

[2] Clark, T. and K. West. 2007. Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics 138, 291-311.

[3] Diether, K. B., C. J. Malloy and A. Scherbina. 2002. Differences of opinion and the cross section of stock returns. Journal of Finance 57, 2113-2141.

[4] Ge, L., T. Lin and N. D. Pearson. 2016. Why does the option to stock volume predict stock returns? Journal of Financial Economics 120, 601-622.

[5] Hu, J. 2014. Does option trading convey stock price information? Journal of Financial Economics 111, 625-645.

[6] Huang, D., F. Jiang, J. Tu and G. Zhou. 2015. Investor sentiment aligned: a powerful predictor of stock returns. Review of Financial Studies 28, 791-837.

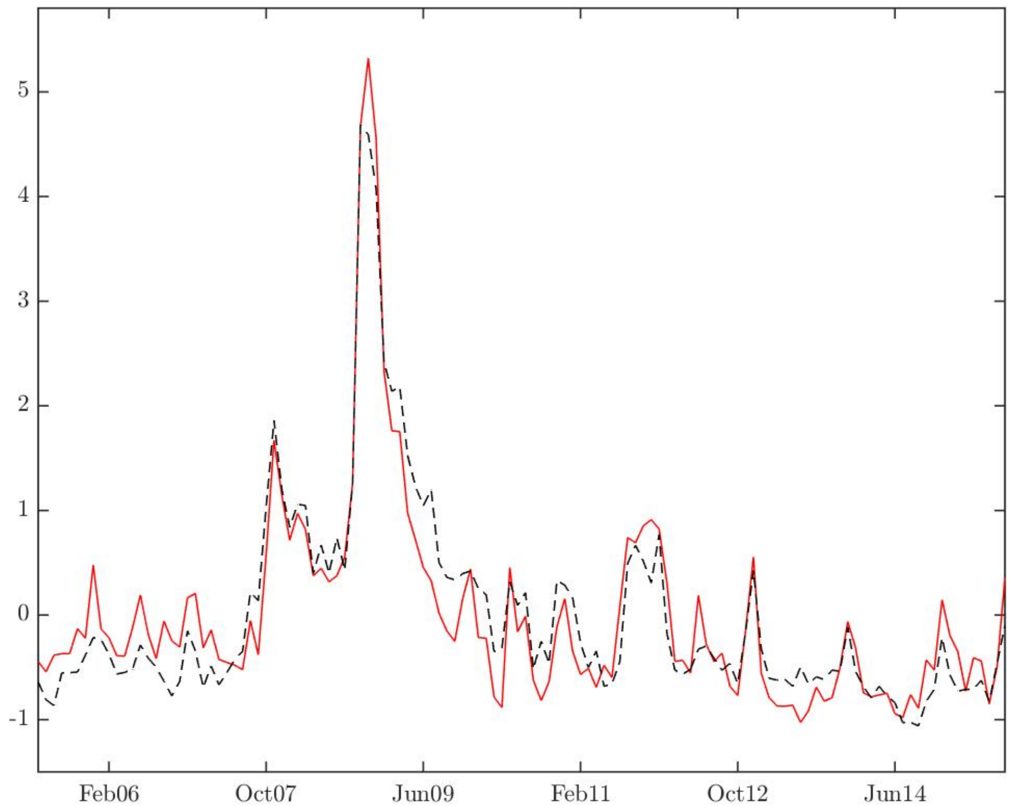
[7] McCracken, M. 2007. Asymptotics for out of sample tests of Granger causality. Journal of Econometrics 140, 719-752.

[8] Newey, W. and K. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703-708.

[9] Neely, C. J., D. E. Rapach, J. Tu and G. Zhou. 2014. Forecasting the equity risk premium: the role of technical indicators. Management Science 60, 1772–1791.

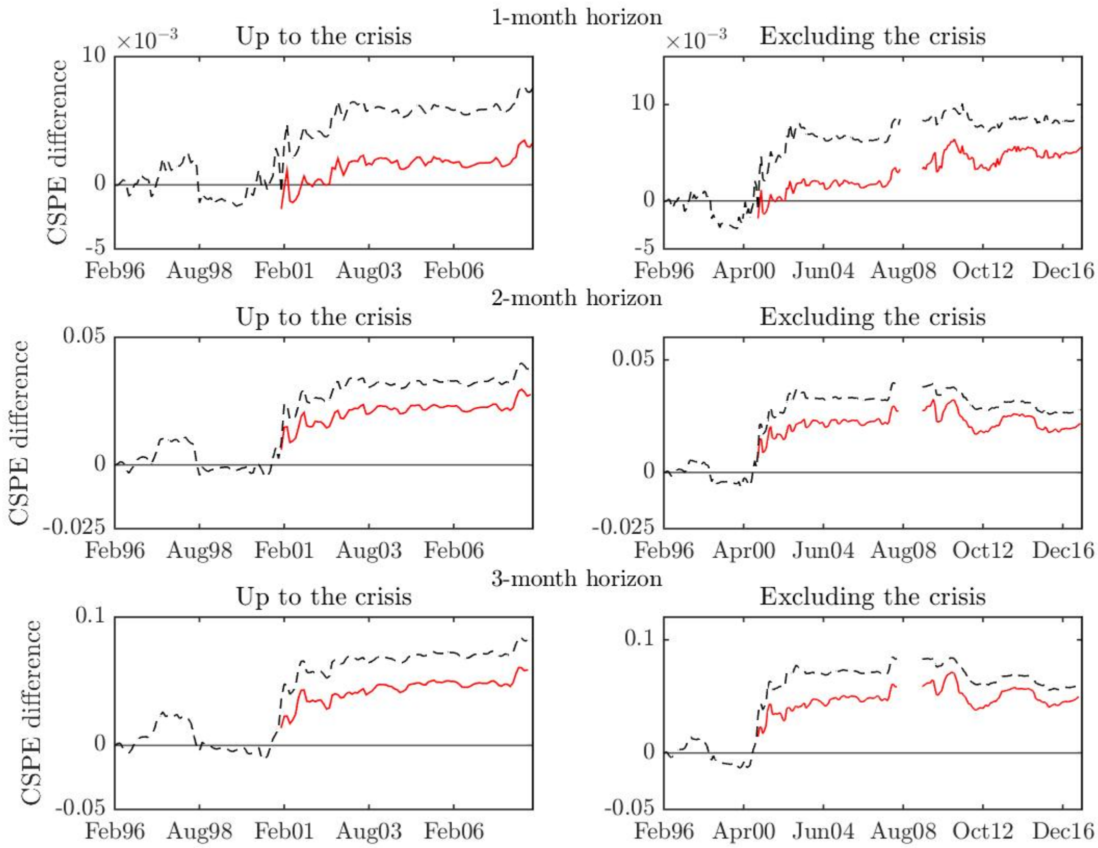
[10] Stambaugh, R. 1999. Predictive regressions. Journal of Financial Economics 54. 375-421.

Figure 1: DISP versus DISP_SIGNED



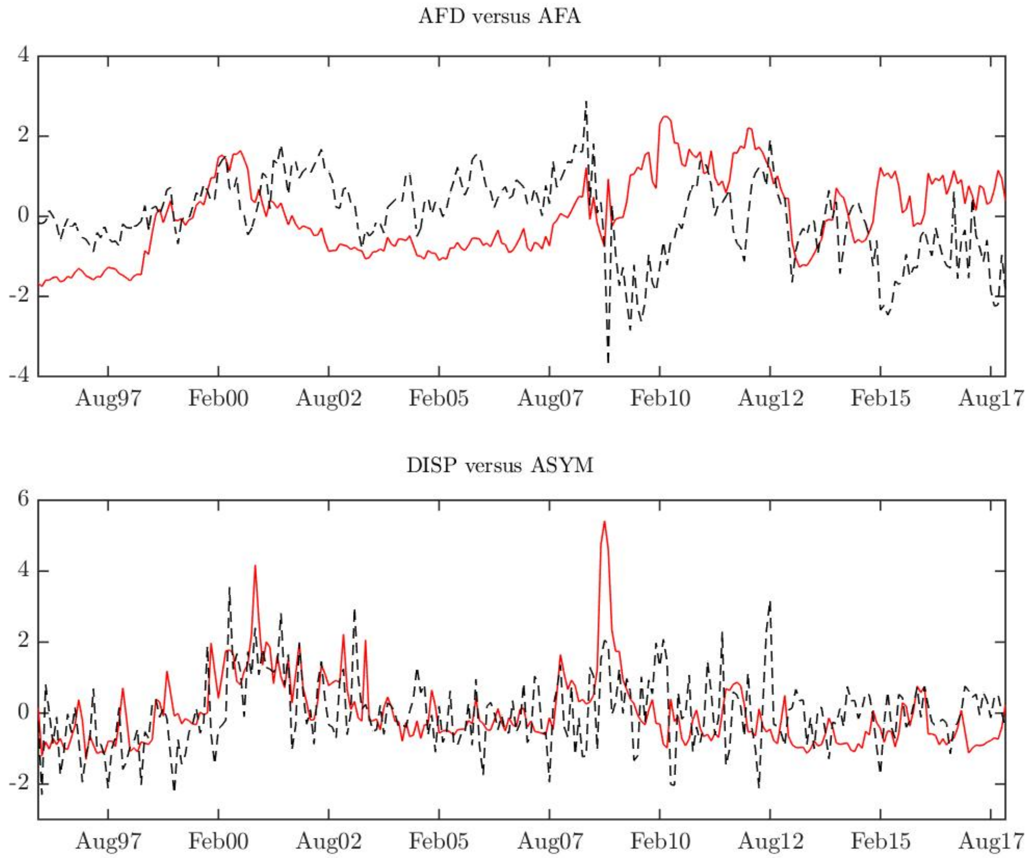
This figure plots the monthly time series of the dispersion in options investors' expectations (DISP, red solid line) versus the dispersion in options investors' expectations with signed volume data (DISP_SIGNED, black dashed line) for the period 2005:05-2015:08. Both variables have been standardized to have zero mean and variance one.

Figure 2: Differences in cumulative squared prediction error for DISP without the financial crisis



This figure plots the differences between the cumulative squared prediction error (CSPE) of the historical average model and the model based on the dispersion in options investors' expectations (DISP). The forecasting horizons are one, two and three months ahead. The black dashed lines correspond to the CSPE from the in-sample analysis, while the red solid lines correspond to the CSPE from the out-of-sample analysis. In the left panels the sample period is 1996:01-2008:06. In the right panels the sample spans the periods 1996:01-2008:06 and 2009:07-2017:12. The out-of-sample period begins in 2001:01 in both cases.

Figure 3: DISP versus ASYM and AFD versus AFA



The top panel of this figure plots the monthly time series of AFD (red solid line) versus AFA (black dashed line). The bottom panel plots the monthly time series of DISP (red solid line) versus ASYM (black dashed line). Both variables have been standardized to have zero mean and variance one. The sample period is 1996:01-2017:12.

Table 1: In-sample predictive power of alternative DISP measures

	1-month horizon		2-month horizon		3-month horizon	
Predictor	$\widehat{\beta}$	R^2 (%)	$\widehat{\beta}$	R^2 (%)	$\widehat{\beta}$	R^2 (%)
DISP_{I}	–10.06	3.47	–8.71	4.66	–8.16	6.06
	(–2.58)**		(–2.91)***		(–2.79)**	
DISP_{II}	–10.14	3.52	–8.49	4.43	–8.09	5.96
	(–2.66)**		(–2.88)**		(–2.75)**	
DISP_{III}	–8.84	2.68	–7.69	3.63	–6.67	4.06
	(–2.35)**		(–2.46)**		(–2.21)**	
DISP_{IV}	–9.91	3.37	–8.63	4.58	–7.96	5.77
	(–2.53)**		(–2.83)**		(–2.66)**	
DISP_{V}	–9.50	3.09	–8.36	4.29	–7.90	5.69
	(–2.37)**		(–2.61)**		(–2.51)**	
DISP_{VI}	–9.18	2.89	–7.42	3.38	–7.44	5.04
	(–2.55)**		(–2.88)***		(–2.97)***	
DISP_{VII}	–10.31	3.64	–8.63	4.58	–7.43	5.02
	(–2.52)**		(–2.66)**		(–2.60)**	
$\text{DISP}_{\text{VIII}}$	–7.50	1.93	–7.15	3.14	–7.25	4.79
	(–1.97)*		(–2.34)**		(–2.52)**	

This table reports the in-sample results for the predictive regressions of the CRSP value-weighted index excess return on alternative dispersion in options investors' expectations (DISP) estimates. DISP_{I} uses the mean absolute deviation, DISP_{II} uses the interquartile range, DISP_{III} uses the standard deviation of normalized strike prices, DISP_{IV} removes options with moneyness between 0.975 and 1.025, DISP_{V} uses options from all available moneyness levels, DISP_{VI} uses only end-of-month values, DISP_{VII} uses only options of the next available standard expiration date and $\text{DISP}_{\text{VIII}}$ uses only options that expire one month after the next available standard expiration date. The sample period is 1996:01-2017:12. Reported coefficients indicate the percentage annualized excess return resulting from a one standard deviation increase in each predictor variable. Newey and West (1987) t-statistics with lag length equal to the forecasting horizon are reported in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% levels based on a wild bootstrap experiment.

Table 2: Out-of-sample predictive power of alternative DISP measures

	DISP_{I}	DISP_{II}	DISP_{III}	DISP_{IV}	DISP_{V}	DISP_{VI}	DISP_{VII}	$\text{DISP}_{\text{VIII}}$
	1-month horizon							
R^2_{OS} (%)	3.04	3.22	2.17	2.66	2.24	2.54	2.97	1.34
MSE-F	6.41**	6.80**	4.52**	5.58**	4.67**	5.31**	6.24**	2.78**
ENC-NEW	7.53**	7.81**	5.72**	7.72**	7.21**	5.64**	7.86**	4.02**
MSE-Adj	1.80**	1.73**	1.58*	1.76**	1.65**	2.04**	1.82**	1.44*
SR	0.99	0.97	0.88	0.99	0.95	0.84	0.98	0.84
ΔCER (%)	7.08	6.73	5.88	7.39	6.99	5.64	7.68	5.53
	2-month horizon							
R^2_{OS} (%)	4.34	4.03	3.09	4.24	3.72	3.63	3.74	2.82
MSE-F	9.17**	8.48**	6.44**	8.94**	7.81**	7.62**	7.86**	5.85**
ENC-NEW	12.22**	11.73**	9.26**	12.37**	11.90**	8.05**	10.95**	7.97**
MSE-Adj	1.64*	1.60*	1.41*	1.59*	1.51*	1.77**	1.66**	1.45*
SR	0.82	0.84	0.69	0.80	0.78	0.59	0.72	0.77
ΔCER (%)	6.31	6.35	4.92	6.30	6.07	3.87	5.84	5.70
	3-month horizon							
R^2_{OS} (%)	5.06	4.81	2.74	4.54	4.10	4.26	3.85	3.28
MSE-F	10.65**	10.11**	5.64**	9.52**	8.55**	8.89**	8.00**	6.78**
ENC-NEW	15.45**	15.65**	9.76**	14.80**	15.10**	12.22**	11.72**	12.20**
MSE-Adj	1.59*	1.51*	1.28	1.53*	1.49*	1.64*	1.64*	1.55*
SR	0.77	0.80	0.63	0.79	0.80	0.53	0.65	0.75
ΔCER (%)	5.91	6.15	4.22	6.12	6.39	3.26	4.85	5.81

This table reports the results of out-of-sample predictability of the CRSP value-weighted index excess return. The total sample period is 1996:01-2017:12 and the forecasting period begins in 2001:01. The forecasting variables are alternative estimates of the dispersion in options investors' expectations (DISP). DISP_{I} uses the mean absolute deviation, DISP_{II} uses the interquartile range, DISP_{III} uses the standard deviation of normalized strike prices, DISP_{IV} removes options with moneyness between 0.975 and 1.025, DISP_{V} uses options from all available moneyness levels, DISP_{VI} uses only end-of-month values, DISP_{VII} uses only options of the next available standard expiration date and $\text{DISP}_{\text{VIII}}$ uses only options that expire one month after the next available standard expiration date. R^2_{OS} is the out-of-sample coefficient of determination, MSE-F is the McCracken (2007) F-statistic, ENC-NEW is the encompassing test of Clark and McCracken (2001) and MSE-Adj is the MSE-Adjusted statistic of Clark and West (2007). ** and * denote significance at the 5% and 10% levels. The critical values for the MSE-F test are 1.518 and 0.616, respectively, while the critical values for the ENC-NEW test are 2.374 and 1.442, respectively. These critical values are based on Monte-Carlo simulations and are provided by the respective studies. We also report the annualized Sharpe ratio (SR) and certainty equivalent return (ΔCER) of a market-timing strategy that is based on each of the predictive models and utilizes mean-variance weights. The benchmark strategy follows the historical average model.

Table 3: Binary market-timing strategies

	DISP	AFD	VRP	TAIL	d-p	e-p	d-e	TERM	DEF	RREL
	Without Short-Sales – 1-month horizon									
SR	0.77	0.37	0.61	0.55	0.53	0.73	0.78	0.39	0.70	0.59
ΔCER (%)	2.50	–1.96	0.74	0.18	0.29	2.50	2.72	–1.70	1.78	0.93
	Without Short-Sales – 2-month horizon									
SR	0.68	0.44	0.61	0.47	0.46	0.64	0.70	0.39	0.72	0.54
ΔCER (%)	2.24	–0.76	1.34	–0.12	–0.28	1.98	2.32	–1.30	2.32	0.75
	Without Short-Sales – 3-month horizon									
SR	0.63	0.34	0.65	0.45	0.50	0.62	0.51	0.38	0.78	0.59
ΔCER (%)	1.73	–2.04	1.84	–0.66	0.24	1.65	0.20	–1.68	2.66	1.31
	With Short-Sales – 1-month horizon									
SR	0.73	0.24	0.53	0.47	0.50	0.72	0.75	0.29	0.65	0.55
ΔCER (%)	5.02	–3.94	1.46	0.35	0.59	5.03	5.47	–3.43	3.56	1.85
	With Short-Sales – 2-month horizon									
SR	0.64	0.32	0.55	0.43	0.42	0.61	0.65	0.31	0.66	0.49
ΔCER (%)	4.51	–1.57	2.67	–0.26	–0.58	3.99	4.66	–2.64	4.65	1.50
	With Short-Sales – 3-month horizon									
SR	0.59	0.21	0.61	0.39	0.48	0.59	0.45	0.29	0.71	0.55
ΔCER (%)	3.48	–4.16	3.71	–1.36	0.49	3.33	0.37	–3.43	5.34	2.63

This table reports the results of market-timing strategies with binary weights based on the out-of-sample predictability of the CRSP value-weighted index excess return. The total sample period is 1996:01-2017:12 and the forecasting period begins in 2001:01. The forecasting variables are the dispersion in options investors' expectations (DISP), analysts' forecasts dispersion (AFD), variance risk premium (VRP), tail risk (TAIL), dividend-price ratio (d-p), earnings-price ratio (e-p), dividend payout ratio (d-e), yield term spread (TERM), default spread (DEF) and relative short-term risk-free rate (RREL). SR stands for the annualized Sharpe ratio and ΔCER is the certainty equivalent return in excess of the historical average strategy.

Table 4: In-sample predictive power of DISP without the financial crisis

Predictor	1-month horizon $\widehat{\beta}$	1-month horizon R^2 (%)	2-month horizon $\widehat{\beta}$	2-month horizon R^2 (%)	3-month horizon $\widehat{\beta}$	3-month horizon R^2 (%)
	Sample period ending before the financial crisis					
DISP	–8.57	2.57	–9.63	6.24	–9.48	9.30
	(–1.85)*		(–2.95)***		(–3.50)***	
	Total sample period excluding the financial crisis					
DISP	–7.09	2.07	–6.38	3.34	–6.30	5.09
	(–2.07)**		(–2.37)**		(–2.64)**	

This table reports the in-sample results for the predictive regressions of the CRSP value-weighted index excess return on the dispersion in options investors' expectations (DISP). In the top panel the sample period is 1996:01-2008:06. In the bottom panel the sample spans the periods 1996:01-2008:06 and 2009:07-2017:12. Reported coefficients indicate the percentage annualized excess return resulting from a one standard deviation increase in each predictor variable. Newey and West (1987) t-statistics with lag length equal to the forecasting horizon are reported in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% levels based on a wild bootstrap experiment.

Table 5: Out-of-sample predictive power of DISP without the financial crisis

	DISP - Sample period ending before the financial crisis	DISP - Total sample period excluding the financial crisis
	1-month horizon	
R^2_{OS} (%)	2.15	1.98
MSE-F	1.98**	3.85**
ENC-NEW	1.95*	4.55**
MSE-Adj	1.50*	2.25**
SR	0.81	1.05
ΔCER (%)	7.79	6.53
	2-month horizon	
R^2_{OS} (%)	8.02	3.73
MSE-F	7.67**	7.28**
ENC-NEW	6.26**	9.04**
MSE-Adj	2.11**	2.04**
SR	0.80	0.89
ΔCER (%)	8.62	5.31
	3-month horizon	
R^2_{OS} (%)	11.64	5.92
MSE-F	11.33**	11.64**
ENC-NEW	9.03**	13.15**
MSE-Adj	2.61**	2.34**
SR	0.74	0.85
ΔCER (%)	7.64	4.71

This table reports the results of out-of-sample predictability of the CRSP value-weighted index excess return. In the left panel the sample period is 1996:01-2008:06. In the right panel the sample spans the periods 1996:01-2008:06 and 2009:07-2017:12. The forecasting period begins in 2001:01 in both cases. The forecasting variable is the dispersion in options investors' expectations (DISP). R^2_{OS} is the out-of-sample coefficient of determination, MSE-F is the McCracken (2007) F-statistic, ENC-NEW is the encompassing test of Clark and McCracken (2001) and MSE-Adj is the MSE-Adjusted statistic of Clark and West (2007). ** and * denote significance at the 5% and 10% levels. The critical values for the MSE-F test are 1.518 and 0.616, respectively, while the critical values for the ENC-NEW test are 2.374 and 1.442, respectively. These critical values are based on Monte-Carlo simulations and are provided by the respective studies. We also report the annualized Sharpe ratio (SR) and certainty equivalent return (ΔCER) of a market-timing strategy that is based on each of the predictive models and utilizes mean-variance weights. The benchmark strategy follows the historical average model.